


```

193 AATGGAAGAAATTTGAGAGAAAGATGTTAGACAAGTCAAAAAGAA 242
509 laalathrglnleualailettyrtyrphethraspserrlaigluasp 555
243 TACCT...CAAAATAAGAAACATTATGACACTTGAAAAAGATTA... 286
526 lysaspplysleuasppltyrhisglsyphesglyaspmetasnaspserth 542
287 AAGGAAAAACAGACAGATTC...TTTGAGCAATTAACATCA... 328
542 rleualaialaalsileuvalglutyrtralaaglnaspsersprop 559
329 ATTGAGGTGAGAAATATTATTATGATAGATGAAAGCAATTAATTAAG 377
559 roclnleuthraspleasphephei1eprasnasnlystygln 575
378 AAGA...GAAATTTGAAATGTTT... 397
576 Serleuilegltthrglntrphisplogluaspplleualaspilleat 592
398 .....AACATTCACACAGAAATGATGCTACTTAAGACAGATTA 438
592 gmetgluaspplysgluvalilleprovalthrhisasnleuthrleua 609
439 AATAGAGAGAGAGAGAT...AGTCAACAGAAAAAATGCAATTAG 482
609 rgltythrvalthrlyleualaglyaspargthrylasphehisphe 625
483 AAAAAAGAAATTTAT...GATCTTAAGAGAGATTTCAACA 520
626 glulleuleuaspasnlysglnleuleuaspserglnthrvally 642
521 GAATTT...AATGAAAAAGAGATGCAATTAACAAAGCTTGA 561
642 straspplysthrasnleuagluheylaspplysalathrileasn 659
562 AAGCAATTAATACAGAAATGAAATAAAGCTGCAATGATGAT 611
659 eulyshtsglyserleuthrleuaglnleuaproglytyrser 675
612 TAAGAGAAATGAGAAACTAAATGATGTACAGATTAAGAAAGAA 661
676 Tyr.....Leuvallysgltthraspserglnlytyrly 687
662 TATGAGAAATGAAATCTTATTAAACAAATCAAAAAAGA... 703
687 svalysvalasnserglnleualalaasnalaithrvalserlythr 704
704 .....CAAGTCATGCACAAAGATCTATGATTAATTCG 737
704 lytlehrserasppluthrleualaphesgluasnlysgluProval 720
738 AAGTGGATTAGATGAAACTGTTGGAAGCTCTAGCAATGAAACAATGAA 787
721 ValProthrGlyValaspGlnlys 728
788 TTACCTATTAAAGCAATGAAAG 811

seq_name: gb_est80:BE919552

```

```

seq_documentation_block:
LOCUS BE919552 626 bp mRNA EST 02-OCT-2000
DEFINITION EST423321 potato leaves and petioles Solanum tuberosum cDNA clone
ACCESSION BE919552
VERSION BE919552.1 GI:10445628
KEYWORDS EST.
SOURCE Solanum tuberosum
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids
I; Solanales; Solanaceae; Solanum.

```

```

REFERENCE 1 (bases 1 to 626)
AUTHORS van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
utterback,T., Hansen,C.L., Doan,B., Boungi,O., Biell,C.R., Ranning
,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
JOURNAL Generation of ESTs from potato leaves and petioles
COMMENT Unpublished (2000)
CONTACT Cathy Ranning
The Institute for Genomic Research
For clone request: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdnaresgen.com.
location/Qualifiers
FEATURES
source
1..626
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTR1D12"
/clone_1lb="potato leaves and petioles"
/tissue_type="leaflets and petioles"
/dev_stage="8 weeks old plants"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI. Tissue was supplied by Dr. Fry (Cornell University).
leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen."
BASE COUNT 172 a 108 c 168 g 178 t
ORIGIN
alignment_scores:
Quality: 108.00 Length: 248
Ratio: 0.878 Gaps: 12
Percent Similarity: 49.597 Percent Identity: 24.194
alignment_block:
US-09-494-297-2 x BE919552 ..
Align seg 1/1 to: BE919552 from: 1 to: 626
503 ThrGlnThrGlnLeuArgAlaAlaThrGlnLeuAlaIleTyrThrPheTh 519
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
9 ACTGAGCTCAGATAAAGACATTTGGAGTGGTATTGGCTTACCATCG 58
519 rAspSerAlaGluLeuaspplysaspplysleuasppl...TyrHisG 534
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
59 AGACTCAACAGGTTTAAGTACGACATTTGGCTGATGCTGGATCCG 108
534 lypheglyaspmetasnaspserthrleualaialaalsileleual 550
||:|||||:|||||:|||||:|||||:|||||:|||||:
109 GAGCAGCTTCACCTGTGAT.....GCAGTTTGTGTATGCGCTGC 149
551 GluTyrAlaGlnaspserasnProGlnLeuThrAspLeuasphe.. 566
|||:|||||:|||||:|||||:|||||:|||||:|||||:
150 GCTTAT.....ATACTTCTTAAGATTCTT 175
567 .....PheIleProasnasnlystyglnSerleuilegltthr 581
|||:|||||:|||||:|||||:|||||:|||||:|||||:
176 GTTTGGTTTCGCTCCACACAGAGAAAGAAATAAGTGGGCTGC 225
581 lntPhisProGluaspLeuValaspIleleargmetgluaspplys 597
||:|||||:|||||:|||||:|||||:|||||:|||||:
226 AGCATCAATCTTGAAGACAGAGATGCGACAGAGAAAG..... 263
598 GluValIleProValThrHisasnleuthrleuargtythrValThrG 614
|||:|||||:|||||:|||||:|||||:|||||:|||||:
264 .....CATTCACGTTCTTCTTCAAGCATTTTCGA 295
614 yLeuAlaGlyaspArgthrylasphe...HispheGluileGluLeu 630
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
296 AGTTGTTAAAGTCGAGCTCGCATGGAATAATGCAATGATGCA 345
630 ysasnasnlysglnleuleuaspserglnthrvallystraspplysThr 646

```

```

346 TCCACCTCTTGCACCTAATCTCGAGATTCATTAG..... 383
      ::::: ||| ::::: |||
647 AsnLeuGlnPheLysAspGlyLysAlaThrIleAsnLeuYshISgLYG1 663
      ::| | | ::::: ||| | | | ::::: ||| | | |
384 GATGCTGACGCAAGTAATTCMAAGCTATTGTGCATCTCATCTTGGGGA 433
      ::| | | ::::: ||| | | | ::::: ||| | | |
663 userLeuThrLeuGlnIleuProGluGlyTyrSerTyrLeuValYsg 680
      | ::| | | | | | | | | | | | | | | | | | | | | |
434 A...ATGAGCTGCAGAGGATAGATGAA...CTGAGTTCTGTGCCAGAG 477
      | ::| | | | | | | | | | | | | | | | | | | | | |
680 IuThrAspSerGluGlyTyrLysValLysAlaSerGlnIuValAla 696
      ||| | | | | | | | | | | | | | | | | | | | | |
478 AA.....ATGCTTACATGATCGAAGTCCA 503
      ::| | | | | | | | | | | | | | | | | | | | | |
697 AsnAlaThrValSerLysThrGlyIleThrSerAspIuThrLeuAlaPh 713
      ::| | | | | | | | | | | | | | | | | | | | | |
504 ACACCT..... 509
      ::| | | | | | | | | | | | | | | | | | | | | |
713 eGluAsnAsnLysGluProValIuProThrGlyValAspGlnLysIleA 730
      ||| | | | | | | | | | | | | | | | | | | | | |
510 .....CCCATATTCTGCTGTGATGTCGAGGTCGCAATAA 543
      | | | | | | | | | | | | | | | | | | | | | | |
730 snGlyTyrLeuAlaLeuIleVal...IleAlaGlyIleSerIeu 743
      ||| | | | | | | | | | | | | | | | | | | | | |
544 ATGCTTGCAATGCAAGGCTGCTGAATTGACAGGTTATCAATT 587
      | | | | | | | | | | | | | | | | | | | | | | |
seq_name: gb_est38:AV603854

```

```

seq_documentation_block: 584 bp mRNA EST 30-AUG-2000
LOCUS AV603854 Bos taurus kidney fetus Bos taurus cDNA clone EIKI019C04
DEFINITION 5', mRNA sequence.
ACCESSION AV603854
VERSION AV603854.1 GI:9734227
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
Bovidae; Bovinae; Bos.
1 (bases 1 to 584)
Sugimoto, Y., Hirotsune, S., Takasuga, A., Itoh, R., Jitohzono, A. and
Suzuki, H.
Bovine cDNA sequencing
Unpublished (2000)
Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics
Odakura, Nishigo, Nishi-Shirakawa, Fukushima 961-8061, Japan
Tel: 81-248-25-5725
Fax: 81-248-25-5725
Email: kazusugie@ccoc.ocn.ne.jp
Single pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.
FEATURES
source
1..584
Location/Qualifiers
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="EIKI019C04"
/clone_lib="Bos taurus kidney fetus"
/tissue_type="kidney"
/dev_stage="fetus"
/lab_host="DH10B"
/note="Vector: pZLI; Site_1: SalI; Site_2: NotI; Poly A
was deleted from a NotI site"

```

```

BASE COUNT 196 a 148 c 131 g 109 t
ORIGIN

```

```

alignment_scores:
Quality: 107.50 Length: 239
Ratio: 1.097 Gaps: 8

```

```

Percent Similarity: 41.004 Percent Identity: 24.268
alignment_block:
US-09-494-297-2 x AV603854 ..
Align seg 1/1 to: AV603854 from: 1 to: 584

```

```

264 AsnLeuSerGlyGlyLeuValProThrLysProProThrProGlyAs 280
      ||| | | | | | | | | | | | | | | | | | | | | |
39 AATTCCTGGAT.....TTACACCCCAATCTCTCCAAACCAAAAC 82
      | | | | | | | | | | | | | | | | | | | | | |
280 pProMepeProPasn.....GlnProGlnThrIh 292
      ||| | | | | | | | | | | | | | | | | | | |
83 CATGTCTTACTCTCCAACTGGAAGACAGCCGAGACCAAGAAAG. 131
      ||| | | | | | | | | | | | | | | | | | | |
292 eValLeuIleArgLysTyrAlaIleGlyAspTyrSerLysLeuGlu 308
      ||| | | | | | | | | | | | | | | | | | | |
132 .....ATTACTACTACATGTGGTCAACAGCAGACCTAGTGGGAT 173
      ||| | | | | | | | | | | | | | | | | | | |
309 GlyAlaThrLeuGlnLeuThrGlyAspAsnValAsnSerPheGlnAlaAr 325
      ||| | | | | | | | | | | | | | | | | | | |
174 CTTCTTACTTGGGAAGCCGAGATGATGCCAGC..... 209
      | | | | | | | | | | | | | | | | | | | | | |
325 gValPheSerSerAsnAspIleGlyGluArgIleGluLeuSerAspGlyT 342
      ||| | | | | | | | | | | | | | | | | | | |
210 .....CTTGACATGAAGCTGACATGAGCTTGGGA 240
      ||| | | | | | | | | | | | | | | | | | | |
342 hTyrThrLeuThrGluLeuAsnSerProAlaGlyTyrSerIleAlaGlu 358
      ||| | | | | | | | | | | | | | | | | | | |
241 CCCACATATGAT..... 254
      ||| | | | | | | | | | | | | | | | | | | |
359 ProIleThrPheLysValGluAlaGlyLysValTyrThrIleIleAspG1 375
      ||| | | | | | | | | | | | | | | | | | | |
254 ..... 254
      ||| | | | | | | | | | | | | | | | | | | |
375 yLysGlnIleGluAsnProAsnLysGluIleValGluProTyrSerValG 392
      ||| | | | | | | | | | | | | | | | | | | |
255 .....GAAATCCCATGAAGACCTCAAAAAGCTTACAGACAG 294
      ||| | | | | | | | | | | | | | | | | | | |
392 IuAlaTyrAsnAspPheGluGluPheSerValLeuThrThrGlnAsnTyr 408
      ||| | | | | | | | | | | | | | | | | | | |
295 AAGCA.....GACACCTCCAGTGAGCGGTAGAAAAGCAA 332
      ||| | | | | | | | | | | | | | | | | | | |
409 AlaLysPheTyrTyrAlaLysAsnLysAsnLysSerGlnValValTyr 425
      ||| | | | | | | | | | | | | | | | | | | |
333 GAAGTATTC.....AGAAAAGAGATGCCAGTTATTCGTCCA 370
      ||| | | | | | | | | | | | | | | | | | | |
425 rCySPheAsnAlaAspLeuLysSerProProAspSerGluAspGlyL 442
      ||| | | | | | | | | | | | | | | | | | | |
371 GTGCTGAACCTTACCGGAAA.....CTGACTGCNAAGTGGGAAGAA 414
      ||| | | | | | | | | | | | | | | | | | | |
442 yStrMetThrProAspPheThrThrGlyGluValLysTyrThrHisIle 458
      ||| | | | | | | | | | | | | | | | | | | |
415 TTACACACACCGAGACTTCAAAACACTAGCTGCAGCTGACATCAGGC 464
      ||| | | | | | | | | | | | | | | | | | | |
459 AlaGlyArgAspLeuPheLysTyrThrValLysProArgAspThrAspPr 475
      ::| | | | | | | | | | | | | | | | | | | | | |
465 GTTATGATATGAAGCTGAAGTACTTGAAGACCCGAGGACCTGGAGTG 514
      ::| | | | | | | | | | | | | | | | | | | | | |
475 oAspThrPheLeuLysHisIleLysValIleGluLysGlyTyrArg 492
      ::| | | | | | | | | | | | | | | | | | | | | |
515 CAATGAGAAATGTGAACACAAACCAAGAGTACATTAGAAAGTACATGC 564
      ::| | | | | | | | | | | | | | | | | | | | | |
492 IuLysGlyGlnAlaIle 497
      ||| | | | | | | | | | | | | | | | | | | |
565 AGAGTTTGGGCGCTGT 581
      ||| | | | | | | | | | | | | | | | | | | |
seq_name: gb_gss23:AZ527959

```

```

seq_documentation_block: 915 bp DNA GSS 03-NOV-2000
LOCUS AZ527959
DEFINITION ENT003117 Entamoeba histolytica sheared DNA Entamoeba histolytica
genomic, DNA sequence.

```

ACCESSION A2527959
 VERSION A2527959.1 GI:11080143
 KEYWORDS GSS
 SOURCE Entamoeba histolytica.
 ORGANISM Entamoeba histolytica
 Entamoeba histolytica
 Eukaryota; Entamoebidae; Entamoeba.
 REFERENCE 1 (bases 1 to 915)
 Loftus, B., Van Aken, S. and Fraser, C.
 AUTHORS Determination of clone end sequences from Entamoeba histolytica
 TITLE HMI:IMS sheared DNA library
 JOURNAL Unpublished (2000)
 COMMENT Contact: Brendan J Loftus
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9713 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: b.loftus@tigr.org
 Clones are derived from the Entamoeba histolytica HMI:IMS sheared
 DNA library
 Seq primer: M13-Forward
 Class: shotgun
 High quality sequence start: 17
 High quality sequence stop: 777.
 Location/Qualifiers
 1..915
 /organism="Entamoeba histolytica"
 /strain="HMI:IMS"
 /db_xref="taxon:5759"
 /clone_lib="Entamoeba histolytica Sheared DNA"
 /note="Vector: pHOSt1; Site_1: Bst I; Constructed at The
 Institute for Genomic Research (TIGR), Rockville, MD.
 Genomic DNA isolated from broth cultures of E. histolytica
 using a method described by Clark and Diamond (Clark,
 C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
 method for isolate identification. Exp. Parasitol.
 77:450.). The DNA was mechanically sheared to give a
 tight size distribution (~2 kb). The v + 1 method used for
 the library construction is described in detail in Smith,
 H.O. and Venter, J.C. (Making small insert libraries for
 whole genome shotgun sequencing projects. In Genome
 Sequencing: A Practical Approach, eds. M. Vaudin and B.
 Barrell, Oxford University Press, 1999)."
 BASE COUNT 429 a 98 c 141 g 247 t
 ORIGIN
 alignment_scores: length: 343
 Quality: 102.50 Gaps: 12
 Ratio: 0.674
 Percent Similarity: 44.315 Percent Identity: 17.784
 alignment_block:
 US-09-494-297-2 x A2527959 ..
 Align seg 1/1 to: A2527959 from: 1 to: 915
 417 LysAsnGlySerGlnValValTyrCysPhe.....AsnA1 429
 ||||| |||
 17 AAAAATTAAATGTCAGAGTAGTAGATTATTTAAATGACAAATG 66
 ||||| |||
 429 aAspLeuLysSerProPaspSerGlnuAspGlyGlyLysThrMetHrp 446
 :|||| :||||
 67 CgTTGCACACTAAACACTAAATTCACAGACAAATTTGAAAAACACA... 112
 :|||| :||||
 446 roAspPheHrHThrGlyGlnValLysTyrThrHisLeaIaGlyArgAsp 462
 :|||| :||||
 112
 463 LeuPheLysTyrThrValLysProArgAspThrAspProAspThrPhele 479
 :|||| :||||
 113AGAGAGAAAATTAAGAAAGATGCTGCTCAACACTTTT 156

479 ULysHisIleLysValIleGluLysGlyTyrArgGluLysGlyGlnA 496
 | :|||||
 157 AACAAAAGAGAGATATATATGATTAAGATGATGAAGAC..... 196
 496 IalIeGluTyrSerGlyLeuThrGluThrGlnLeuArgAlaIaIaThrGln 512
 | :|||||
 196 196
 513 LeuAlaIeTyrThrPheThrAspSerAlaGluLeuAspLysAspLysIle 529
 | :|||||
 197TGTTTTACTGAGAGATGATTAATGATCAAAAAAGACT 234
 : : :
 529 ULysAspTyrHisGlyPheGlyAspMetAsnAspSerThrLeuAlaVal. 545
 : : :
 235 AATTAATATTAAGAAATAGATTAAGATAAAAAAGAAATTAAGATTAAGATA 284
 : : :
 546 ..AlaLysIleLeuValGluTyrAlaGlnAspSerAsnProGlnLeu 561
 | | |
 285 ATGATAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 331
 : : :
 562 ThrAspLeuAsp.....PheP 567
 : : :
 332 AGTGATGTTGAAAAAATAGTTCAAACATTCAGATTCAGATTAATTTCTA 381
 : : :
 567 eIleProAsnAsnLysTyrGlnSerIleuIleGlyThrGlnTrpHisP 584
 : : :
 382 TACACTGACAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 431
 : : :
 584 roGluAspLeuVal.....AspIleLeuArgMetGluAspLysLysGlu 598
 | | |
 432 TAGAAGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 481
 : : :
 599 ValIleProValThrHisAsnLeuThrLeuArgLysThrValThrGlyLe 615
 | : :
 482 AATATACCAAAATCAATGATATACATGCAATGCAATGTTGAATGAA 531
 : : :
 615 uAlaGlyAspArgThrLysAspPheHisPheGluIleGlnLeuLysAsnA 632
 : : :
 532 ACCAAAGATTAATA...AAGCAATTCACAGAAATATTTCTTTCAGATC 578
 : : :
 632 snLysGlnGluLeuLeuSerGlnThrValLysThrAspLysThrAsnLeu 648
 : : :
 579 TTGAAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 601
 : : :
 649 GluPheLysAspGlyLysAlaThrIleAsnLeuLysHisGlyLysSerIle 665
 : : :
 601 601
 665 uThrLeuGlnGlyLeuProGluGlyTyrSerTyrLeuValLysGluThra 682
 : : :
 602AAGAAATTCACAAAGACTTCTTT.....GAAG 632
 : : :
 682 spSerGluGlyTyrLysValLysValAsnSerGlnGluValAlaAsnAla 698
 | | |
 633 ATCAAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 670
 : : :
 699 ThrValSerLysThrGlyIleThrSerAspGluThrLeuAlaPheGlnuAs 715
 : : :
 671 TCATTAATAACTAATCATCTATTTTGCACGATTAATTAATTAATTAATAAAC 720
 : : :
 721 AGAAGTACTGATACCTCTCTATATATTAATTAAGTACTTAAAGATG 770
 : : :
 727 LnuLysIleAsnGlyTyrLeuAlaIleuIle 736
 : : :
 771 AGAAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 799
 : : :
 seq_name: gb_est49:AW654454
 seq_documentation_block: 508 bp mRNA EST 14-JUL-2000
 LOCUS AW654454
 DEFINITION 103924 MARC IBOV Bos taurus cDNA 5', mRNA sequence.

```

ACCESSION      AM654454
VERSION        AM654454.1
KEYWORDS       EST.
SOURCE         GI:7420280
ORGANISM       Bos taurus
               cow.
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
               Bovidae; Bovine; Bos.
REFERENCE      1 (bases 1 to 508)
AUTHORS        Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M.,
               Bennett,G.A., Fahrnkung,S.C., Freking,B.A., Rohrer,G.A., Laegreid
               ,W.W. and Keeler,J.W.
               Design and use of four pooled tissue normalized cDNA libraries for
               EST discovery in cattle
JOURNAL        Unpublished (2000)
COMMENT        Contact: Smith TPL
               USDA, ARS, US Meat Animal Research Center
               PO Box 166, Clay Center, NE 68933-0166, USA
               Tel: 402 762 4366
               Fax: 402 762 4390
               Email: smithte@mail.marc.usda.gov
               Single pass sequencing. Bases called and alt_trimmed with phred
               v0.980904.e. Vector identified by cross_match with the -mismore 18
               and -mismatch 12 options.
FEATURES       PCR Primers
               FORWARD: AGGAAACAGCTATGACCAT
               BACKWARD: GTTTTCCAGTCACGACG
               Plate: 88 row: B column: 3
               Seq primer: ATTTAGTGACACTATAG.
               location/Qualifiers
               1..508
               /organism="Bos taurus"
               /db_xref="taxon:9913"
               /clone_lib="MARC 180Y"
               /russue_type="pooled"
               /lab_host="DH10B"
               /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
               library made from pooled tissue from lymph node, ovary,
               fat, hypothalamus, and pituitary."
BASE COUNT    169 a 117 c 135 g 87 t
ORIGIN
alignment_scores:
               Quality: 101.50 Length: 129
               Ratio: 1.538 Gaps: 4
               Percent Similarity: 51.163 Percent Identity: 31.008
alignment_block:
US-09-494-297-2 x AM654454 ..
Align seg 1/1 to: AM654454 from: 1 to: 508
379 GUAaNPProASnlySGluIleValGluProTYrSerValGluAlaTYrAs 395
|||||  |||  ::|||  ::::| ||||
111 GAAATTCCTCATGAGACCTCAAAAAAGCCTAAGACAGCAGCAAGCA..... 155
395 nAspHeGluGluPheserValLeuThrThrGlnAsnTYrAlaLysPhet 412
::: ||| |||::: ::::: |||
156 .....GACACCTCCAGCTGAGCTGGCTAAGAAACCAAGAAATTC. 197
412 YTYrAlaLysAnlysaAnGlySerSerGluValAlaTYrCysPheAsn 428
:::|:::| |:::| |:::| |:::| |:::|
198 .....AGAAAGAGATGTCCTCAATTATGCTCCAGGCGCTGAAG 236
429 AlaAsPLeuLysErProAsPserGluAspGlyGlyLysThrMetH 445
237 CCTTACCGGAAA.....CCTGACTGCCAAAGTGGGAAGATTAACCAAC 280
445 rProAsPHeThrThrGlyGluValAlaLysTYrThrHisIleAlaGlyArg 462
| ||||| | ||| ||| ||||| |:::
281 CGAAGACTTCAAAACACTAGCTCGCAAGCTGACTCTACGGCGTTATATGAATA 330

```

```

462 spleuPhylystYrThVallysprOArGspHrAspPArOAspPAr 478
    ::|||  |||  |||  |||  |||  |||  |||  |||  |||  |||
479 LeuTSHSLetLySValleGluDySGLYrArGsluSgLygl 495
    ::|||  |||  |||  |||  |||  |||  |||  |||  |||  |||
381 GCGAAACACCAAAACCAAGAGACTACATTAAGAGACACTGCAATG 430
    431 GCGCTGTT...TACAAACCCAAAGAGACACTGAGTTA 464

seq_name: gb_est28:AL037842

seq_documentation_block:
LOCUS      AL037842          570 bp.          mRNA          EST          29-FEB-2000
DEFINITION DKEZP564H077..t1 564 (synonym: hfbz2) Homo sapiens CDNA clone
ACCESSION  DKEZP564H077 5', mRNA sequence.
VERSION    AL037842
KEYWORDS   AL037842.1  GI:5407173
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 570)
AUTHORS   Bloecker,H., Boecker,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and
            Wiemann,S.
TITLE      EST (Bloecker, et al.)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Bloecker H
            MPS
            Am KlopferSpitz 18a D-82152 Martinsried, Germany
            This is the 5' sequence of the clone insert
            clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
            sequenced by GGF (National Research Centre for Biotechnology Ltd.,
            Braunschweig/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No sl sequence available.
            This clone (DKEZP564H077) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, Germany; Email: clone@rzpd.de.
            Location/Qualifiers
                source
                    1. 570
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="DKEZP564H077"
                        /clone_1lb="564 (synonym: hfbz2)"
                        /tissue_type="brain"
                        /dev_stage="fetal"
                        /lab_host="X1-2blue"
                        /note="Vector: PAMP1, Site_1: NotI, Site_2: SalI"
BASE COUNT      190 a      118 c      152 g      110 t
ORIGIN
alignment_scores:
    Quality: 101.50      Length: 129
    Ratio: 1.538      Gaps: 4
    Percent Similarity: 51.163      Percent Identity: 31.008

alignment_block:
US-09-494-297-2 x AL037842 ..

Align seg 1/1 to: AL037842 from: 1 to: 570

379 gGuAsnPrOAsnLySGluILleValGluPrOtySerValGluAlaTyraS 395
|||||  |||  ::|||  ::|||  |||  |||  |||  |||  |||  |||
130 GAAACCCCAATGAGGCTCGAAAAAGCCAAAGCCAGACAGACAAACA..... 174

395 nAspHeGluGluPrSerValLeuThrThGlnAsnTyraLalysPheT 412
:::  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

175GACACCTCCAGTACAGCAAGAAAGCAAGAGTATTC. 216
 412 YTYTAlAlAlAsnLysAsnGlySerSerGlnValValTYrCysPheAsn 428
 217AGAAAGAGATGCCAGTTCATCCTCAGCGCTGAC 255
 429 AlAspLeuLysSerProProAspSerGlnAspGlyGlyLysThMetTh 445
 256 CCTACCGGAAA.....CTGACTGCAGAAAGTGGCAAGATTACACACAC 299
 445 rProAspPheThrThrGlyGluValLysTYrThHisIleAlaGlyArgA 462
 300 TGAAGACTTAAACATCTGCTCCGACCTGATCACCCTGTATGAAATA 349
 462 sPLeuPheLysTYrThrValLysProArgAspThrAspProAspThrPhe 478
 350 AGGACCTGAAGTACTGTAGAAATCTGAGACCTGGACTGCATGAGAT 399
 479 LeuLysHisIleLysValIleGluLysGlyTYrArgGluLysGlyG 495
 400 GTGAACACCAAAACAGAGTACATTAAGAAATGATGACGAATTTGG 449
 495 nAlaIleGluTYrSerGlyLeuThrGluThrGlnLeu 507
 450 GGCTGTT...TACAAACCCAAAGAGACACCTGAATTA 483
 seq_name: gb_est28:AL037979

seq_documentation_block: 673 bp mRNA EST 29-FEB-2000
 LOCUS AL037979 DKFZ564P077_r1 564 (synonym: hfr2) Homo sapiens cDNA clone
 DEFINITION DKFZ564P077_5', mRNA sequence.
 ACCESSION AL037979
 VERSION AL037979.1 GI:5407291
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 673)
 Bloeker, H., Boecker, M., Brandt, P., Mewes, H.W., Gassenhuber, J. and
 Wiemann, S.
 TITLE EST (Bloeker, et al.)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Bloeker H
 MIPs
 Am Kioferspitz 18a D-82152 Martinsried, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by GBF (National Research Centre for Biotechnology Ltd.,
 Braunschweig/Germany) within the cDNA sequencing consortium of the
 German Genome Project.
 SI sequence also available.
 This clone (DKFZ564P077) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
 location/Qualifiers

FEATURES
 source
 1..673
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="DKFZ564P077"
 /clone_lib="564 (synonym: hfr2)"
 /tissue_type="brain"
 /dev_stage="fetal"
 /lab_host="X1-2blue"
 /note="Vector: PAMP1; Site_1: NotI; Site_2: SalI"

BASE COUNT 222 a 147 c 171 g 133 t
 ORIGIN

alignment_scores: Quality: 101.50 Length: 129

Ratio: 1.538 Gaps: 4
 Percent Similarity: 51.163 Percent Identity: 31.008
 alignment_block:
 US-09-494-297-2 x AL037979 ..

Align seg 1/1 to: AL037979 from: 1 to: 673

379 GlnAsnProAsnLysGluIleValGluProTYrSerValGluAlaTYrAs 395
 |||||
 130 GAAGACCCATGAGGCGCTCGAAAAAGCCAGAGACGAGAGCA..... 174
 395 nAspPheGluLupPheSerValLeuThrThrGlnAsnTYrAlaLysPhe 412
 175GACACCTCCAGTACAGCAAGAAAGCAAGAGATATTC. 216
 412 YTYTAlAlAlAsnLysAsnGlySerSerGlnValValTYrCysPheAsn 428
 217AGAAAGAGATGCCAGTTCATCCTCAGCGCTGAC 255
 429 AlAspLeuLysSerProProAspSerGlnAspGlyGlyLysThMetTh 445
 256 CCTACCGGAAA.....CTGACTGCAGAAAGTGGCAAGATTACACACAC 299
 445 rProAspPheThrThrGlyGluValLysTYrThHisIleAlaGlyArgA 462
 300 TGAAGACTTAAACATCTGCTCCGACCTGATCACCCTGTATGAAATA 349
 462 sPLeuPheLysTYrThrValLysProArgAspThrAspProAspThrPhe 478
 350 AGGACCTGAAGTACTGTAGAAATCTGAGACCTGGACTGCATGAGAT 399
 479 LeuLysHisIleLysValIleGluLysGlyTYrArgGluLysGlyG 495
 400 GTGAACACCAAAACAGAGTACATTAAGAAATGATGACGAATTTGG 449
 495 nAlaIleGluTYrSerGlyLeuThrGluThrGlnLeu 507
 450 GGCTGTT...TACAAACCCAAAGAGACACCTGAATTA 483
 seq_name: gb_est78:BE744259

seq_documentation_block: 797 bp mRNA EST 15-SEP-2000
 LOCUS BE744259 601576452F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3837373 5',
 DEFINITION mRNA sequence.
 ACCESSION BE744259
 VERSION BE744259.1 GI:10158251
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 797)
 NIH-MGC http://mgc.nci.nih.gov/.
 TITLE NIH-MGC
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert_Strausberg@nih.gov
 Tissue Procurement: DCTD/DRP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: Image.llnl.gov
 plate: LLCW522 row: h column: 14
 High quality sequence stop: 744.
 location/Qualifiers

FEATURES
 source
 1..797
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3837373"

```

/clone.lib="NIH_MGC.9"
/tissue.type="adenocarcinoma cell line"
/lab.host="DH10B (phage-resistant)"
/notes="Organ: ovary; Vector: pORF7; Site.1: XhoI; Site.2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT 241 a 180 c 197 g 176 t 3 others

ORIGIN

alignment_scores:

Quality:	101.50	Length:	129
Ratio:	1.538	Gaps:	4
Percent Similarity:	51.163	Percent Identity:	31.008

alignment_block:

US-09-494-297-2 x BE744259 ..

Align seg 1/1 to: BE744259 from: 1 to: 797

```

379 GLuSnPrOaSnLySGuIlleValGluProTySeRValGluAlaTyRAs 395
|||||
30 GAAAAACCCCATGAAGGCTCGAAGAAAGCCCAAGACAGACAGCA.... 74
395 nAsPheGluGluPheSeRValLeuThrThrGlnAsnTyRAlaLysPheT 412
75 .....GACACCTCCAGTGAACCTAGCAAGAAAGCAAGAAAGATATTC. 116
412 yRTyRAlaLysAsnLysAsnGlySeRSeRValValTyRcysPheAsn 428
117 .....AGAAAGAGATGTCCAGTTCATCGTCACGTGCTGCTGAC 155
429 AlaSpLeuLysSeRProProAsPSeRGlusPlyGlyLysThrMetH 445
156 CCTTACCGGAAA.....CCTGACTGCAGAGTGGAGAAATACCAAC 199
445 rProAsPheThrThrGlyGluValLysTyRThrHisIleAlaGlyArg 462
200 TGAAGACTTTAAACATCTGCTCGCAAGCTGCTACGCGTGTATGAAT 249
462 sPluLeuPheLysTyRThrValLysProArgAsPThrAsPProAsP 478
250 AGAGCTGAAGTACTGTAGAAATCCTGAGGACGTGAGTGCATGAGAT 299
479 LeuLysHisIleLysLysValIleGluLysGlyTyRArgGluLysG 495
300 GTGAACACCAAAACCAAGAGTACATTAAAGATACATGACAGAAATT 349
495 nAlaIleGluTySeRgLyLeuThrGluThrGlnLeu 507
350 GGTCTGT...TACAACCCCAAGAGACACTGAATTA 383
```

seq_name: gb_est37:AV400861

seq_documentation_block:

LOCUS AV400861 796 bp mRNA EST 05-FEB-2000

DEFINITION AV400861 Bombyx mori brain Daizo P0 (just after pupation) Bombyx

ACCESSION AV400861

VERSION AV400861.1 GI:6904513

KEYWORDS EST.

SOURCE domestic silkworm.

ORGANISM Bombyx mori

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea; Bombycidae; Bombyx.

REFERENCE 1 (bases 1 to 796)

AUTHORS Mita,K., Morimoto,M., Shimada,T., Okano,K. and Maeda,S.

TITLE Bombyx mori cDNA

JOURNAL Unpublished (2000)

CONTACT: Mita K

COMMENT Genome Research Group National Institute of Radiological Sciences Anagawa 4-9-1, Inage, Chiba 263-8555, Japan Email: kmitsa@irs.go.jp method:uni-directional, sequence direction:sequenced from T3 primer (5' -> 3')

Project="Silkworm Genome Program in MAFR and Research for the Future Program in JSPS". see "Silkbase", <http://www.ab.a.u-tokyo.ac.jp/silkbase/>, for whole ESTdb.

FEATURES

Location/Qualifiers

source 1..796

/organism="Bombyx mori"

/strain="Daizo"

/db_xref="taxon:7091"

/clone="Br-1899"

/clone.lib="Bombyx mori brain Daizo P0 (just after pupation)"

/sex="female/male mixed"

/tissue_type="brain"

/dev_stage="P0 (just after pupation)"

BASE COUNT 332 a 123 c 174 g 167 t

ORIGIN

alignment_scores:

Quality:	101.00	Length:	244
Ratio:	0.835 <td>Gaps:</td> <td>10 </td>	Gaps:	10
Percent Similarity:	49.590 <td>Percent Identity:</td> <td>20.902 </td>	Percent Identity:	20.902

alignment_block:

US-09-494-297-2 x AV400861 ..

Align seg 1/1 to: AV400861 from: 1 to: 796

```

473 ThrAsPProAsPThrPheLeuLysHisIleLysLysValIleGluLysG 489
|||||
16 ACTGAACACTGATTCCTTAATTAATTAACAAATTT..... 54
489 yTYRArgGluLysGlyGlnAlaIleGluTySeRgLyLeuThrGluThrG 506
55 .....CGAGCTGTTTACAAATGCTCAACGGGCAACGA 88
506 InLeuArg AlaAlaThrGlnLeuAlaIleTyRTyRThrAsPSeRAl 522
89 AGATCAAAACGCTGGAATTCGATGCT.....AAACCTCTGACGAGGA 132
522 aGluLeuAsPlyAsPlyLysLysAsPtyRHisGlyPheGlyAsPHeT 539
133 GGAAGAGAGTATTTTACAAACACATACGAGGAGGTTTATGATGTTG 182
539 sNAsPSeRThrLeuAlaValAlaLysIleLeuValGluTyRAlaGlnAs 555
183 AAGAGGAC.....GTAGATATCAACACAGAA 208
556 SeRAsnProGluLeuThrAsPLeuAsPhePheIleProAsnAsnAs 572
209 AAGGAAGTTGATGATGAAGTTCATGATTTGATATGATGAGAAATGA 258
572 n.....LysTyRGlusSerLeuIleG 579
259 TGAACCAAGTGTCTGACAGCAAGAAATTGAAGAAAGAAAGAAAGTTG 308
579 lYThrGlnTrpHisProGlu..... 585
309 GCACCTAAAGCTTTATAGAGCCCAATTAAGAAAGAAATTAATGAATCA 358
586 .....AsPLeuValAsPIL 590
359 AAAAAGACAGTCATCAAAAATCCAAAGATGATTAATCTATAGAGACAA 408
```



```

682 GCATCATGACGAGCGAGCGTGCACCTCCAGAAAGTTATGATAGGTAC 731
656 .....Thrlleasnleuylshisgllyserleuthleuglnlyl 670
732 AAGAGTCACAGGCGCTTATGAAATGTGGAAAGCATCGAAGATTAAGAG 781
670 euprogiluglytyrserlyrleuvalylsleuthraspsersgllytyr 686
782 AGCCGGAAGATGTTCTGTGACCGGCTCAGGACCTGAACAGGCGGATG 831
687 LysVallylsValasnSerGlnGluValAlaAsnAlaThrValSerlysth 703
832 GCGAACGCGGAGAGAAACCAAGGAGGAGAACGAGATTAAGCGGACAC 881
703 rgllylIethSeraspIuThrleuAlaPheGluAsnAsnlysluProw 720
882 ACGGCTCCAGGAGGAGCTGGAATAATGCGCAACAGAGAGAGCGCG 931
720 alValProThr 723
932 CAATCCCAACC 942

```

seq_name: gb_est30:AU139155

seq_documentation_block:

LOCUS AU139155 800 bp mRNA EST 25-OCT-2000
DEFINITION AU139155 PLACE1 Homo sapiens cDNA clone PLACE1010047 5', mRNA
sequence.

ACCESSION AU139155
VERSION AU139155.1 GI:11000676
KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 800)
Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,Y., Makamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.

TITLE HRT human cDNA project

JOURNAL Unpublished (2000)

COMMENT Contact: Takao Isogai
Genomics Laboratory

Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3951
Fax: 81-438-52-3952

Email: genomics@hri.co.jp

HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix
Research Institute; cDNA library construction; Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

FEATURES Location/Qualifiers

source 1..800
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="PLACE1010047"
/clone_lib="PLACE1"
/rname="PLACE1"
/rsize_type="PLACE1"
/note="Vector: pME18SFL3"

BASE COUNT 212 a 227 c 184 g 173 t 4 others
ORIGIN

alignment_scores:
Quality: 99.00 Length: 272
Ratio: 0.780 Gaps: 16
Percent Similarity: 46.691 Percent Identity: 24.632

alignment_block:

US-09-494-297-2 x AU139155

Align seg 1/1 to: AU139155 from: 1 to: 800

```

503 ThrcgluthrGlnleuArgAlaAlaThrGlnleuAlaIleTytrTyPheTh 519
70 ACCATCCACAGTTTACACACGACGACGTACACAGATCTACCTGTACAC 119
519 rAspSerAlaGlnleuAspLysAspLysleuLysAspTyrlshisgllyrPheG 536
120 C..... 120
536 LysMetLsnAspSerThrleuAlaValAlaLysIleleuValGluTy 552
121 .....TTGAATGACAAATGCTCGAGC...TCCCTGTGGTCATCGAGCC 162
553 AlaglnAspSerAsnProproglnleuThrAspLsnAspPhePheIleP 569
163 TCCACTGCCATGTATGACCA...TCCACTGCGGTTTCCGTGGCCAC 206
569 AsnAsnAsnLysTyrlGlnSerleuIleGlyThrGlnTyrlshisPro.... 584
207 CACACCCCAAT.....TCCTTGGCTG...GTATCATGCGACCGGCCAC 244
585 ....GluAspLeuValAspIleIleArgMetGlu.....Asp 595
245 GTGCGCAGATTTACCGGCTCATCATCATCAATGATGAGAACCTGGGCTCT 294
596 LysLysGlnValIleProValIThrHisAsnleuThrleuArgLysThVa 612
295 CCAGAGAAAGTGTCTCTCGGCCCCCGGCTGTGTCTCACAGAGCTACTAT 344
612 lThrGlyleuAlaGlyAspArgThrLysAspPheHisPheGluIleGlu 629
345 TACTGGCCTGGAA...CCGGGAACCGAATATACATTTATGTATGATGCC 391
629 euLysAsnAsnLys..... 633
392 TGAAGATAATTCAGAACGAGCGACCCCTGATTGGAAAGAAAGACAGGA 441
634 GlnGlnleuLeuSerGlnThr..... 640
442 CAGAGACTCTCTCTCAGACACACATCTCATGGGCCCATTCAGAGCAC 491
641 .....ValLysThrAspLysThrAsnL 648
492 TTTCGACTACATCATTTTCATGTCATCTGTGGCACTGATGAAGAACCT 541
648 euGlnPheLys...AspGlyLysAlaThrIleAsnleuLysIshisglly 663
542 TACAGTTCAGGCTTCCTCGAATCTTACC..... 570
664 SerleuThrleuGlnGlyleuProGlnGlyTyrlSerTyrlleuVally 679
571 AGTGCACCTGTGACAGGCTTCACACAGGTGCCACCTACACATCATCAAT 620
679 sgluthraspsersgllytyrlyslsValylasnSerGlnGluVala 696
621 GAGGCGCCTGAAGACGACGAGGATAGGTTCGGGAGAGAGTTGTGA 700
696 laasnAlaThrValSerLys.....ThrglylelthSeraspGl 709
671 CCGTGGCAACTCTGTCAACGAGGCTTGAACCAACCACTACGAGGATCTC 720
709 uThrleuAlaPheGluAsnAsnLysgluProwalProThrGlyVala 726
721 G.....TGCTTGACCCCTACACACAGTTTCATTTATGCCGTTTGAGATG 764
726 spGlnLysIleasn 730
765 ATGGGGAANCAAT 778

```

seq_name: gb_gss23:A2535645

seq_documentation_block:

LOCUS A2535645 875 bp DNA GSS 03-NOV-2000
 DEFINITION ENTOM497F Entamoeba histolytica Sheared DNA Entamoeba histolytica
 genomic DNA sequence.
 ACCESSION A2535645
 VERSION A2535645.1 GI:11092592
 KEYWORDS GSS.
 SOURCE Entamoeba histolytica.
 ORGANISM Entamoeba histolytica.
 Eukaryota; Entamoebidae; Entamoeba.
 REFERENCE 1 (bases 1 to 875)
 Loftus,B., Van Aken,S. and Fraser,C.
 Determination of clone end sequences from Entamoeba histolytica
 HMI:IMSS sheared DNA library
 Unpublished (2000)
 JOURNAL Contact: Brendan J Loftus
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: b.loftus@eigf.org
 Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
 DNA library
 Class: shotgun
 Seq primer: M13-Forward
 High quality sequence start: 22
 High quality sequence stop: 749.
 Location/Qualifiers
 1..875
 /organism="Entamoeba histolytica"
 /strain="HMI:IMSS"
 /db_xref="taxon:5759"
 /clone_lib="Entamoeba histolytica Sheared DNA"
 /note="Vector: PHOS1; Site 1: Bst I; Constructed at The
 Institute for Genomic Research (TIGR), Rockville, MD.
 Genomic DNA isolated from broth cultures of E. histolytica
 using a method described by Clark and Diamond (Clark,
 C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
 method for isolate identification. Exp. Parasitol.
 77:450.). The DNA was mechanically sheared to give a
 tight size distribution (~2 kb). The v + 1 method used for
 the library construction is described in detail in Smith,
 H.O. and Venter, J.C. (Making small insert libraries for
 whole genome shotgun sequencing projects. In Genome
 Sequencing: A Practical Approach, eds. M. Vaudin and B.
 Bartell, Oxford University Press, 1999)."

BASE COUNT 516 a 80 c 144 g 135 t
 ORIGIN

alignment_scores:
 Quality: 99.00 Length: 279
 Ratio: 0.723 Gaps: 16
 Percent Similarity: 49.104 Percent Identity: 21.864

alignment_block:
 US-09-494-297-2 x A2535645 ..

Align seg 1/1 to: A2535645 from: 1 to: 875

483 LysLysValIleGluLys.....GlyTyrArgLysGluGlnAla 496
 ||||||||||||||||
 29 AAAAAAGGTAATGTGAAGACGAAAGACGTTAAAGAAAAACAAAGA 78
 |||||
 496 aileglutyrserglyleuthrglnleuargalalathrglnl 513
 ||| :|||
 79 TTGGGAG.....GAAAAGACATTAAAGCTATA..... 106
 ||| :|||
 513 euallalterytyrphethraspseralacluleuaspLysaspLysleu 529
 :||| :|||
 107GAAGCAAGAAAT 118
 :||| :|||
 530 LysaspTyrHisglYpHeGlYaspMetAsnspSerThrleuAlaValAl 546

||||: : : : :
 119 AAAAAACAAAACCAATTAAGAAATGATGAT...AGATGTAATAATGC 165
 :|||: : : : :
 546 aLysIleleuValGluTyrAlaGlnaspSerAsnProGlnleuThra 563
 :|||: : : : :
 166 AAGAAATAATA..... 175
 563 spleuAspPhePheIleProAsnAsnLysTyrGlnserleuIleGly 579
 ||||| :|||
 176AATAACAAA...CAAGCTAATATGAA 199
 :|||: : : : :
 580 ThrGlnTrpHisProGluAspLeuValAspIleleuArgMetGlu...As 595
 :|||: : : : :
 200 ACACAAAGAAAGAAACAGAAATTAAGAAAGAAAGCTAGACAGAAAGA 249
 :|||: : : : :
 595 pLysLysGluValIleProValThrHisAsnleuThr...leuArgLysT 611
 :|||: : : : :
 250 TAAAGAAAGAAATCTGCTGCAAAATCAAAATTAAGTGAATTAAGAAA 299
 :|||: : : : :
 611 hValThrGlyleuAlaGlyAspArg..... 619
 :|||: : : : :
 300 AACTTAAGAAATTCAGCATGGAAGAAATGGAAGAAAGAACCAATCTCA 349
 :|||: : : : :
 620ThrlYsaspPheHisph 625
 :|||: : : : :
 350 GAAGCAATGATGAGAAAGCAAAAGAAATAGTTAGGAAGCAATTTGAGA 399
 :|||: : : : :
 625 egluile.....GluLeuLysAsnLysG 634
 :|||: : : : :
 400 AGAAATTAACAGAAATGAACACAAATTAAGAAATCAACAGAAATTAAT 449
 :|||: : : : :
 634 InGluLeuSerGlnThrValLysThrAspLysThrAsnleuGluPhe 650
 :|||: : : : :
 450 AAGCAAGAAAGAAAGAAATGAAGAAATTAAGAAAGAAAGAAATTAAG 499
 :|||: : : : :
 651 LysaspLysLysAlaThrIleAsnleuLysHisglYluserleuThr... 666
 :|||: : : : :
 500 CAA...GCAAAA.....ATTAATGATAGTAAAGAAAGAAATACACAGA 540
 :|||: : : : :
 667LeuGlnGlyleuProGluGlyTyrSert 676
 :|||: : : : :
 541 AGTAATAATAAAGAAAGAAATCCGACGATCCAGCTAATTAAGAA 590
 :|||: : : : :
 676 YrleuValLysGluThrAspSerGluGlyTyrLysValLysAsnSer 692
 :|||: : : : :
 591 AATTTGAAAAAGAA...AATGCAAGAAATGAAGAAAGAAATTTGAAGA 634
 :|||: : : : :
 693 GlnGluValAlaAsnAlaThrValSerLysThrGlyIleThrSerAsp.. 708
 :|||: : : : :
 635 CTAAAAAAGAGAAACAAACGCTAGAACAAATGCAACACACTAGAGAA 684
 :|||: : : : :
 709GluThrleuAlaPheGluAsnAsnLysGlu 718
 :|||: : : : :
 685 GAAATTTGAACGATAGAG...GAAATTAATAAGAA 718
 :|||: : : : :
 seq_name: gb_est81:BE970296

seq_documentation_block:
 LOCUS BE970296 862 bp mRNA 04-OCT-2000
 DEFINITION 601679453R1 NIH_MGC_78 Homo sapiens cDNA clone IMAGE:3949908 3',
 mRNA sequence.
 ACCESSION BE970296
 VERSION BE970296.1 GI:10583229
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 862)
 NIH-MGC http://mgc.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: CLONETECH Laboratories, Inc.
 cDNA Library Preparation: CLONETECH Laboratories, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLC815 row: 1 column: 13.
 Location/Qualifiers

FEATURES

source

```

1..862
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3949908"
/clone_1ib="N1H_MGC_78"
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: pancreas; Vector: pDNR-LIB (Clontech);
Site_1: Sfil (99ccctcgcc); Site_2: Sfil (99ccatagcc
); 5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-ATTCTAGAGCGCGAGCGCGGACATG-dt(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.2 kb (range 0.5-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."

```

BASE COUNT 147 a 243 c 256 g 216 t

ORIGIN

alignment_scores:
 Quality: 98.50 Length: 233
 Ratio: 0.912 Gaps: 11
 Percent Similarity: 46.352 Percent Identity: 22.747

alignment_block:

US-09-494-297-2 x BE970296/rev ..

Align seg 1/1 to reverse of: BE970296 from: 1 to: 862

```

112 LysLysAlaPheProLeu.....GlySerAspSerSerVa 123
:::|||||:::|||||
610 CAAAGAGCTTCTCCAGAGCGCGCCCAACGAGCGCGCCCAAGGGGT 561
123 LysLysTrpTrpLysHisAspGlyIleSerThrLys..... 136
||||| ||||| |||||
560 CACCAATGATGATAAAATCCCGGGGGAATAACACAGAGCGCGCC 511
||||| ||||| |||||
137 .....PheGluAspTyrAlaMetSerProArgIleThrGlyAsp 149
||| ::| ::| ::|
510 CCAGGATTTTCTTCCAGACAGCGCGGGAAGCC..... 476
||| ::| ::| ::|
150 GluLeuAsnGlnLysLeuArgAlaValMetTyrAsnGlyHisProGlnAs 166
||||| ::| ::| ::|
475 .....CAGAGTTGGCGCCCGCAGATACAAACAGCGCGCCAA 435
||||| ::| ::| ::|
166 nAlaAsnGlyIleMetGluGluLeuGluProLeuAsnAlaIleArgValT 183
||| ||| |||
434 C.....CTTGTCAGATTAAAGGCTA 415
|||||
183 hncGlnGluAla...ValTrpTyrTyrSerAspAsnAlaProIleSerAsn 198
|||||
414 CACGAAATACAGACAGTTCTCTCCGGAATATAGAGCGCCGAC 365
::| |||||
199 ProAspGluSerPheLysArgGluSerGluSerAsnLeuValSerThrSe 215
::| |||||
364 CAGAGAGCGAGTTTC.....ACGAGCCCTTTCGAACCC 330
|||||
215 rGlnLeuSerLeuMetArgGlnAlaLeuLysGlnLeuIleAspPro.... 230
||| |||||
329 GGATTGCGACATCATGCCGAGGGGAT.....CTGGCGCAACATCAAA 286
|||||

```

```

231 .....AsnLeuAlaThrLysMetPro 237
|||||
285 AGAATGTGGAGAACCAACACAGGTAATAAGAGAGATCAACCA 236
|||||
238 LysGlnVal.ProAspAspPheGlnLeuSerIlePheGluSerGluAspL 254
|||||
235 AGGAGATATCCCCC.....CGGATCTCCATGGCAGAGAGAGAG 198
|||||
254 ysgLysAspLysTyr.....AsnLysGlyTyrGlnAsnLeuLeu 266
|||||
197 GGACCCAGCAATTTCTAGTGGCGCCCAACCATGGGAGACCATCTCCAG 148
|||||
267 SerGlyGlyLeuValProThrLysProProThrProGlyAspProPhe 283
|||||
147 AATGACAGATACGGCGTCCGAGAACCCCGAGGGGAGAAATGCTGCG 98
|||||
283 tProProAsnGlnProGlnThrThrSerValLeuIleArgLysTyrAla 300
|||||
97 TCCCGCTCCCTACCCGCTGTGGAGACATCTATTGGCAGCAGACGGG 48
|||||
300 LegLysAspTyrSerLysLeuLeuGluGlyAlaThrLeuGlnLeuThr 315
|||||
47 CAGGCTTCTTCAAAAACAAACAAAGGGGATACCCACAAATGACA 1
|||||

```

seq_name: gb_est26:AI895477

seq_documentation_block: 566 bp mRNA

LOCUS AI895477 EST 27-JUL-1999

DEFINITION EST264920 tomato callus, TMU Lycopersicon esculentum cDNA clone

ACCESION CLEC7N6, mRNA sequence.

VERSION AI895477 GI:5601379

KEYWORDS EST.

SOURCE tomato.

ORGANISM Lycopersicon esculentum

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids

I; Solanales; Solanales; Solanum; Lycopersicon.

1 (bases 1 to 566)

REFERENCE Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.

, Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ann,S., Roning

, C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.

Generation of ESTs from tomato callus tissue

Unpublished (1999)

CONTACT: David Frisch

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 4366

Fax: 864 656 4293

Email: dfisch@clemson.edu

5 prime sequence.

FEATURES

source

1..566 Location/Qualifiers

/organism="Lycopersicon esculentum"

/cultivar="TA496"

/db_xref="taxon:4081"

/clone="CLEC7N6"

/lab_host="tomato callus, TMU"

/tissue_type="callus"

/dev_stage="25-40 days old"

/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:

XhoI; supplier: Giovannoni laboratory; cLEC - cotyledons

of seedlings 7-10 days post-germination were excised, cut

at both ends and placed on MS medium with no selection.

Mixed callus was harvested at 25 and 40 days and included

undifferentiated masses. Tomato Callus EST Library"

BASE COUNT 192 a 96 c 137 g 141 t

ORIGIN

alignment_scores:

Quality:	98.00	Length:	176
Ratio:	1.101	Gaps:	9
Percent Similarity:	50.568	Percent Identity:	27.841

alignment_block:
US-09-494-297-2 x AI895477

Align seg 1/1 to: AI895477 from: 1 to: 566

131 AspGluIleSerThrLysPheGluAspLysPyrAlaMetSerProArgIleTh 147
132 GATGGAGATCTTAAGTCACTATTTCACACATTAAGTGGCATTAATGCCAAC 201
147 rGlyAspGluLeuAsnGlnLysLeuArgAlaValMetLysAsnGlyHisP 164
148 TGAATGATGAGCCCTCAACACCTCTCTATTCATTCTGAAGCAAAATGCA... 247
164 rGlnAlaAsnAlaAsnGlyIleMetGluGlyLeuGluProLeuAsnAlaIle 180
248ATTGAT.....GTCAACCTCG... 256
181 ArgValThrGlnGluAlaValTrpTyrTyrSerAspAsnAlaProIleSe 197
257ATTGAT.....GATGAGCGTCAAGATATA 273
197 rAsnPro.....AspGluSerPheLys.....ArgGluSerGluS 209
274 AAACCCCAAGACATGATGATGAGAGCTATCAAGAGGAAAGCATATCTGAGA 323
209 eArgAsnLeuValSerThrSerGlnLeuSerLeuMetArgGlnAlaLeuLys 225
324 GTAAC.....TCTGCGCTGATCTTTTCAACTATATACAGAAATCCCTTCAA 367
226 GlnLeuIleAspProAsnLeuAlaThrLysMetProLysGlnValProAs 242
368ATTGAT.....TACCGAAGAGAAGAGAAA 393
242 pAspPheGlnIleuSerIlePheGluSerGlnAspLysGlyAspLysTyr 259
394 AGAC.....TTGTCAGATTAAGTATTAATCCACAAAGCTGAAAAGGCAA 437
259 snLysGlyTyrGlnAsnLeuLeuSerGlyGlyLeuValProThrLysPro 275
438 GTAACAACTATTAAGAGAACACTCTCTGGGGTGGCTT.....GAGAGC 478
276 ProThrProGlyAspProProMetProProAsnGlnProGlnThrIleSe 292
479 CCGAGAGCTGAGGTCCCAAAACCTCAATACACTGCAAAAGATTAAATTCAAA 528
292 rValLeuIleArgLysTyrAlaIleGly 301
329 GAAGTATTAACAAGTCAATATCAATTGGGT 556

